Strategic Management Journal
Strat. Mgmt. J., **31**: 390–412 (2010)
Published online EarlyView in Wiley InterScience (www.interscience.wiley.com) DOI: 10.1002/smj.816
Received 25 August 2006; Final revision received 17 September 2009

INDUSTRY LEARNING ENVIRONMENTS AND THE HETEROGENEITY OF FIRM PERFORMANCE†

NATARAJAN BALASUBRAMANIAN[1] and MARVIN B. LIEBERMAN[2]*
[1] *College of Business Administration, Florida International University, Miami, Florida, U.S.A.*
[2] *UCLA Anderson School of Management, Los Angeles, California, U.S.A.*

This paper characterizes interindustry heterogeneity in rates of learning-by-doing, and examines how industry learning rates are connected with firm performance. Using plant-level data from the U.S. manufacturing sector, we measure the industry learning rate as the coefficient on cumulative output in a production function. We find that learning rates vary considerably among industries and are higher in industries with greater R&D, advertising, and capital intensity. More importantly, we find that higher rates of learning are associated with wider dispersion of Tobin's q and profitability among firms in the industry. These findings suggest that learning intensity represents an important characteristic of the industry environment that affects the range of firm performance.

INTRODUCTION

Industries vary considerably in the degree to which firm performance is determined by learning from direct operating experience or learning-by-doing. In some industries, products and processes may be relatively simple, or entrepreneurs and managers may be able to leverage external sources (e.g., specialized technology suppliers, consultants, or competitors' employees) to acquire knowledge about their business operations. Other industry environments may not support such acquisition of knowledge or may involve complex, knowledge-intensive processes and products, thereby constraining firms to improve performance largely through direct experience. In such environments, learning-by-doing may significantly affect firm performance.

In this study, we focus on the importance of accumulated experience in the production process as a measure of the importance of learning-by-doing in an industry ('industry learning intensity'). Further, we examine how differences in industry learning intensity are associated with business performance. Using plant-level data from the U.S. Census Bureau (USCB) on over 55,000 manufacturing plants during the time period 1973 to 2000, we estimate the industry learning rate as the coefficient on prior cumulative output in a production function. Applying these industry learning rates to firm data from Compustat, we find that the cross-sectional variation in business performance within an industry, as measured by the interpercentile range (10th to 90th) of firm q and firm profitability, is much greater in industries

Keywords: learning; firm heterogeneity; RBV; productivity
* Correspondence to: Marvin B. Lieberman, UCLA Anderson School of Management, 110 Westwood Plaza, B415, Los Angeles, CA 90095-1481, U.S.A.
E-mail: marvin.lieberman@anderson.ucla.edu
† DISCLAIMER: The research in this paper was conducted while the authors were Census Bureau research associates at the California Census Research Data Center (CCRDC). Research results and conclusions expressed are those of the authors and do not necessarily indicate concurrence by the Bureau of the Census. The results presented in this paper have been screened to ensure that no confidential data are revealed.

with higher learning intensity. These findings suggest that learning intensity is an important characteristic of the industry environment that should be considered in studies of firm and industry performance.

This study draws from literature on organizational learning and 'learning curves,' which have been studied since the 1930s. The 'learning curve'—the empirical relationship between unit cost of production and operating experience—has been estimated for numerous industries such as ships (Rapping, 1965; Thornton and Thompson, 2001), chemicals (Lieberman, 1984), and semiconductors (Gruber, 1994; Gruber, 2000). Cost reductions generally appear to follow a 'power-law,' that is, the unit cost of production declines at a decreasing rate with increasing experience, typically measured as prior cumulative output. While most studies have found that performance improves as organizations accumulate operating experience, the rate of learning has been shown to vary greatly across industries. In a review of 22 field studies on learning-by-doing, Dutton and Thomas (1984) noted that unit costs fell at rates ranging up to 45 percent for each doubling of cumulative experience. Moreover, learning rates have been found to vary within an industry—even within subunits of the same firm. In their examination of productivity, Hayes and Clark (1986) found that learning rates differed significantly even across factories within the same company. In an analysis of cardiac surgery departments implementing a new technology for minimally invasive cardiac surgery, Pisano, Bohmer, and Edmondson (2001) found that the learning curve slope varied significantly across organizations. While such studies have demonstrated that learning rates vary among organizations and industries, prior investigations have drawn from limited datasets and have not attempted to characterize differences in learning rates across a broad range of industries.

This study also links to another line of papers, mostly in the structure-conduct-performance literature, that examines how industry factors affect firm performance. Many empirical studies have examined how variables such as industry structure, research and development (R&D) intensity, and advertising intensity affect firm performance (see Schmalensee [1989] for a review.). However, neither the industry learning intensity nor the role of direct experience has been studied (empirically) as a variable that could affect firm performance. This is a bit surprising given that a number of studies have argued that the learning curve has implications for competitive strategy and may be used to generate 'first mover advantages' (e.g. Spence, 1981; Lieberman, 1987).

This study makes two contributions to the existing literature on learning. First, it provides a broad-brush characterization of plant-level learning-by-doing in over 100 three-digit standard industrial classification (SIC) code industries in the U.S. manufacturing sector. This characterization reinforces findings in prior studies that industries vary considerably in their learning rates. In addition, we provide a set of reasonably comparable industry-level estimates of the importance of learning from direct experience. Most prior studies have focused on a single product or service, largely due to nonavailability of longitudinal data across industries. In this study, we use a large sample drawn from USCB data that spans the entire U.S. manufacturing sector. We adopt a production function approach and measure the industry learning intensity as the coefficient on prior cumulative output in a production function. This approach is approximately equivalent to the traditional unit-cost learning curve and provides a reasonably uniform measure of learning rates across industries, albeit subject to some limitations. We find that the industry learning rate displays considerable heterogeneity across industries and that it is positively correlated with industry capital-labor ratio, R&D intensity, and advertising intensity, even after controlling for joint industry-year fixed effects or plant fixed effects. These correlations are consistent with the intuitive notion that learning-by-doing may be more important in industries with greater complexity.

Second, our study demonstrates that industry learning intensity has robust relationships with firm performance. In particular, we find that the cross-sectional heterogeneity of firm performance within an industry, as measured by the interpercentile range of firm profits or firm q, is higher in industries with higher rates of learning. In other words, in such industries, the difference between the 'best' and the 'worst' (conditional on survival) firms is considerably higher. Taken together, our findings add to the existing literature by introducing industry learning intensity as an important component of the industry environment that may explain competitive heterogeneity.

DOI: 10.1002/smj

LEARNING-BY-DOING

Learning-by-doing is generally considered to be the result of organizational search for better routines combined with trial and error experimentation, though it has been modeled in a number of different ways (Levitt and March, 1988; Muth, 1986; Jovanovic and Nyarko, 1995). In this paper, we use the information-theoretic model developed by Jovanovic and Nyarko (1995) to develop hypotheses that characterize interindustry variations in the rate of learning and relate the rate of learning-by-doing to heterogeneity of firm performance.[1] This model not only relates characteristics of the underlying learning processes to the learning rate, but also allows us to examine the impact of changes in the learning rate on the heterogeneity of firm performance.

Here, we summarize the model briefly. (Technical details can be found in the Jovanovic and Nyarko [1995] paper.) Decision makers (e.g., managers, engineers, workers) make decisions that affect the efficiency of a production activity. The efficiency is determined by how far the production decisions are from their 'ideal' values. More specifically, the efficiency η is defined as:

$$\eta = \Theta\ \Pi[1 - (y_j - z_j)^2]_{j=1 \text{ to } N} \tag{1}$$

where N is the number of tasks that activity requires, z_j is the decision for the j^{th} task, and y_j is the 'ideal' for the j^{th} task.[2] Note that efficiency is maximized at z=y, and the maximal level of efficiency is Θ. The ideal level 'y' is a random variable that the decision makers do not have complete information about, prior to production. Specifically, it is assumed that

$$y = \theta + w \tag{2}$$

where θ represents the optimal way (on average) to perform the activity, and w represents transitory disturbances that have zero mean and variance σ^2_w. Decision makers know the variance of θ, σ^2_θ, but do not know its mean. Based on information available before a production run, decision makers choose z for that run. Upon completing the production run, decision makers observe the resulting efficiency η, and use that 'signal' to revise their estimates of the mean of θ. As the number of production runs increases, decision makers have increasingly precise estimates of the mean of θ, but they never know it exactly because of the presence of disturbances.

This formulation, though simple, incorporates three distinct dimensions of complexity. The first, N, is the most intuitive. The greater the number of tasks that any production activity requires, the greater the number of decisions involved, and hence, the higher the complexity of the activity. However, activities that entail a large number of tasks need not be highly complex. It is possible that even though the number of tasks involved is large, the decision makers have a lot of information about how the tasks should be performed, and hence the uncertainty surrounding the optimal decision is small. The variance of θ, σ^2_θ, is the second dimension of complexity, which captures the uncertainty about the optimal way to perform a specific task. Interpreted this way, tasks that are relatively new to industry participants are likely to have a greater variance. The third distinct dimension is the importance of transitory disturbances w (as measured by the variance, σ^2_w). In situations with low levels of such disturbances or 'noise,' decision makers can glean more useful information from each production run than they can in contexts where these disturbances are high.

A fourth dimension of complexity, ignored by the model, is the degree of interaction among the tasks. Interactions can greatly increase system complexity (Simon, 1962), and in extreme cases, learning can become so difficult that little or no progress takes place (Levinthal, 1997). In the Jovanovic and Nyarko (1995) model, tasks are performed sequentially and there are no interactions. Most industrial manufacturing processes are made up of sequential tasks that broadly fit the model assumptions. It is possible, however, that some industries are characterized by a degree of interactive complexity sufficient to negate the predictions drawn below.

We argue that the three dimensions of complexity in the Jovanovic and Nyarko (1995) model capture important elements of the learning environment in manufacturing plants, and moreover, that these elements are likely to vary greatly across

[1] We thank an anonymous referee for pointing us in this direction.

[2] The implications do not depend on the choice of this particular functional form. Jovanovic and Nyarko (1995) show that the findings from their model are robust to different functional forms.

DOI: 10.1002/smj

industries. For example, one would expect complexity along all three dimensions to be low in mature industries where technology is well understood, and where the production process consists of a small number of stages that can be readily observed. Representative examples in our sample include leather goods and yarn production. At the other extreme, complexity is likely to be high in industries such as computer manufacturing and petroleum refining, which incorporate uncertainty at a large number of process steps. Given the rapid pace of change in computer technology, knowledge of optimal methods from previous product generations provides only limited guidance for current practice, so uncertainty is high in a new plant. In the case of petroleum refining, work-in-process can be monitored only indirectly, and variations in the quality of crude oil can make it hard to identify the optimal process parameters. As these examples suggest, at a broad level one can conceive of the complexity of manufacturing plants as an increasing function of the number of process stages (N), and the degree of uncertainty (σ^2_θ) and 'noise' (σ^2_w) arising at each stage.

Based on this formulation, Jovanovic and Nyarko (1995) derive a formula for the expected efficiency on production run τ,

$$E_\tau(\eta_\tau) = \Theta\ (1 - x_\tau - \sigma^2_w)^N, \qquad (3)$$

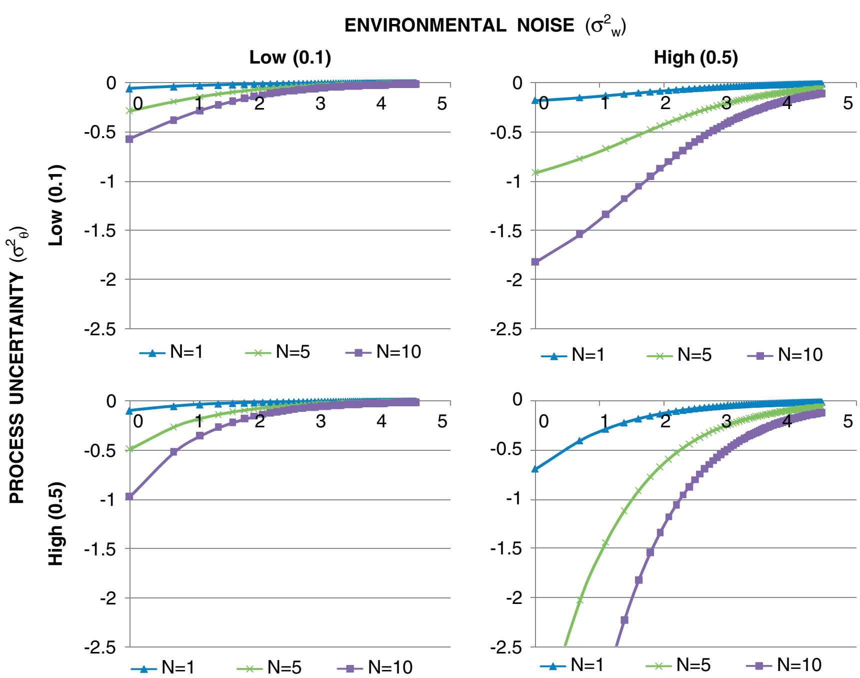
where $x_\tau = \sigma^2_w \sigma^2_\theta / (\sigma^2_w + \tau\sigma^2_\theta)$. Noting that $x_\tau = 0$ as the number of production runs tends to infinity, we can define the eventual expected efficiency as

$$E(\eta^*) = \Theta\ (1 - \sigma^2_w)^N. \qquad (4)$$

Dividing Equation 3 by Equation 4, we obtain a learning curve that is a function of the three dimensions of complexity.

$$\rho = (1 - x_\tau - \sigma^2_w)^N / (1 - \sigma^2_w)^N. \qquad (5)$$

Equation 5 allows us to analyze the relationship between the dimensions of complexity and the slope of the learning curve. We do this in Figure 1 where we plot the logarithm of this learning curve for different values of the parameters. It is evident



The horizontal axis in these graphs is the logarithm of the cumulative number of production runs. The vertical axis is the logarithm of the ratio of expected efficiency at any point in time to the eventual expected efficiency.

Figure 1. Learning rate and complexity. This figure is available in color online at www.interscience.wiley.com/journal/smj

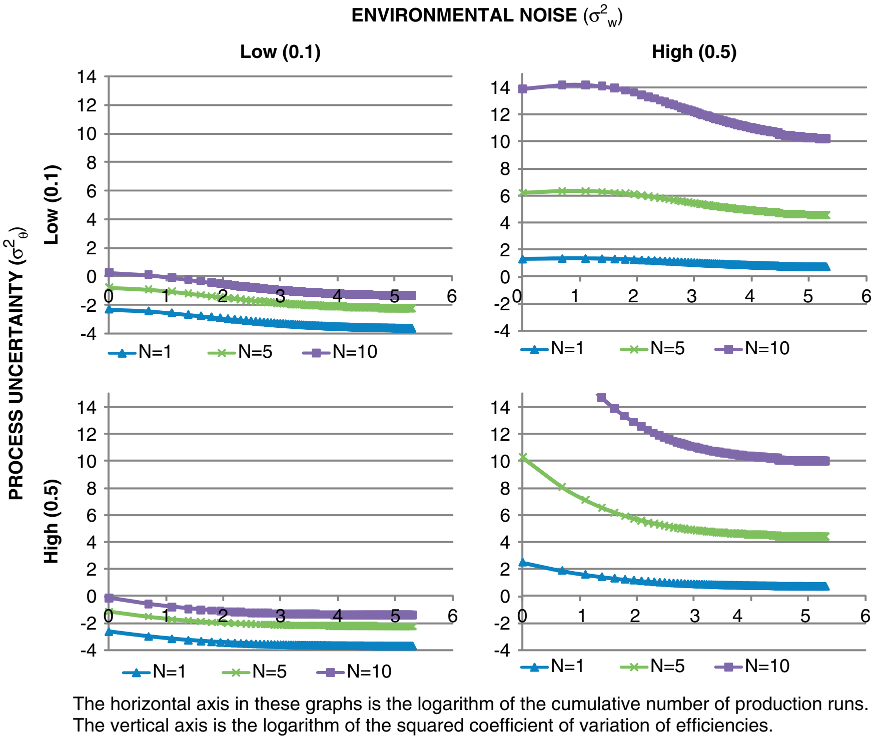


Figure 2. Learning rates and performance heterogeneity. This figure is available in color online at www.interscience.wiley.com/journal/smj

that as complexity becomes greater, as measured by any of the three dimensions, the slope of the learning curve increases. The underlying intuition is that in more complex situations, decision makers have to sift through more data to collect the same amount of information on the optimal way to organize production. Hence, they tend to start at a lower level of efficiency (relative to the maximum) and require more production experience to reach the potential maximum. These arguments lead to our first hypothesis.

Hypothesis 1: The rate of learning-by-doing, as measured by the slope of the learning curve, will be higher in industries with greater complexity.

The model also provides a testable hypothesis regarding the relationship between industry learning intensity and heterogeneity of performance. Jovanovic and Nyarko (1995) define the heterogeneity of performance in terms of the squared coefficient of variation of efficiency (i.e., the variance divided by the square of the mean), $\nu(\eta_\tau)$, and derive an expression for this measure of heterogeneity

$$\nu(\eta_\tau) = [1 + \mathrm{Var}(\eta_\tau)/[\mathrm{E}_\tau(\eta_\tau)]^2]^{\mathrm{N}-1} - 1 \qquad (6)$$

where $\mathrm{Var}(\eta_\tau) = 2\sigma^4{}_{\mathrm{w}}[1 + \tau.\sigma^2{}_\theta/(\sigma^2{}_{\mathrm{w}} + \tau.\,\sigma^2{}_\theta)^2]$ and $\mathrm{E}_\tau(\eta_\tau)$ is as in Equation 3.

Figure 2 plots the logarithm of the measure of heterogeneity in Equation 6 versus the logarithm of the cumulated number of production runs. The relative inequality for any cohort of firms is high initially and then eventually decreases to an asymptotic level.[3] The intuition behind this pattern is that even though the decision makers may all start with the same level of experience, some receive more favorable signals than others—either due to luck or ability—which generates inequality among the decision makers. As time progresses, the impact of this difference decreases, although it never becomes zero because of the presence of transitory disturbances. More interestingly, the

[3] Also, as observed by Jovanovic and Nyarko (1995), for some parameter values, the relative inequality increases initially before it starts decreasing.

DOI: 10.1002/smj

heterogeneity is positively associated with complexity. Irrespective of the dimension of complexity, Figure 2 shows that as complexity becomes greater, the heterogeneity of observed efficiencies increases. The intuition is the same as before—the difference in the quality of signals received by the decision makers leads to differences in the observed efficiencies. As complexity rises, the variance of the signals increases, thereby increasing the observed heterogeneity in efficiency.

Equation 6 and the plots in Figure 2 describe the heterogeneity in efficiency within cohorts having the same number of production runs, τ. Hence, a point on one of the graphs in Figure 2 represents the relative inequality among firms with the *same* production experience. However, the total heterogeneity within an industry is not only a function of variation within cohorts, but also of differences across cohorts of varying production experience. While the model does not provide an explicit formula, we can infer that intercohort variation will increase with the slope of the learning curve. (Given two cohorts with different levels of production experience, a steeper learning curve implies that the difference in efficiency between these cohorts will be greater.) Figure 1 shows that the slope of the learning curve increases with the three dimensions of complexity in the model. Thus, it follows that intercohort heterogeneity will also increase in these dimensions of complexity.

This leads us to our second hypothesis.

Hypothesis 2: The heterogeneity of firm performance will be greater in industries with higher rates of learning.

MEASURING INTENSITY OF LEARNING-BY-DOING

The traditional approach to measuring learning-by-doing for a product is to estimate a power-law function of the following form:

$$C = AX^{-\lambda} \tag{7}$$

where C is the unit cost of the product; A is a constant; X is a measure of experience, typically prior cumulative production; and $\lambda > 0$ is the rate of learning-by-doing.

This formulation is purely empirical and is a reduced form representation of the various processes of learning from direct experience. The disadvantage of this approach is that it requires detailed cost and production quantity data, which are not easily available for a large number of firms. Our method for measuring learning-by-doing follows Bahk and Gort (1993) and is a variant of the traditional approach. Bahk and Gort (1993) incorporate learning-by-doing within a production function and estimate the coefficients using data from individual manufacturing plants. Following this approach, we can write:

$$Y_{ijt} = \Phi_{jt}(K_{ijt})^{\alpha}{}_{j}(L_{ijt})^{\beta}{}_{j}(X_{ijt})^{\lambda}{}_{j}\nu_{ijt} \tag{8}$$

where Y is the current period real value added, measured as real revenues less real materials expenses; Φ is a constant (explained below); K and L are real capital stock and quantity of labor, respectively; X is prior cumulative output, a measure of experience; α, β, and λ are all positive and less than 1; ν is a plant-specific term (explained below); and subscripts i, j, and t refer to plant 'i,' industry 'j,' and year 't,' respectively.

This formulation is an extension of the widely used Cobb-Douglas production function. In addition to the usual inputs of capital and labor, prior operating experience is considered an 'input' into the production process in the sense that a higher level of operating experience increases output for any given level of capital and labor. Hence, λ, the coefficient on prior experience, denotes the industry learning intensity.

We can interpret the learning coefficients obtained from this approach in two ways. First, the coefficient λ can be interpreted in a straightforward manner as the importance of learning (from direct experience) in the production process. A higher value of λ implies a greater role for accumulated experience in the production process. We could also interpret learning to be an improvement in 'productivity' resulting from experience. Productivity (or more precisely, total factor productivity) as defined in the economics literature is a measure of the efficiency of *physical* resource use. Hence, firms with higher productivity have the capability to generate more or better quality output using the same amount of physical resources. There are two physical resources considered above, capital and labor. So, we could define productivity of a plant as $P_{ijt} = \Phi_{jt}(X_{ijt})^{\lambda}\nu_{ijt}$, which is simply the right hand side of Equation 8 excluding the inputs of physical resources. The second term of this expression, ${X_{ijt}}^{\lambda}$, is the increase in productivity resulting

DOI: 10.1002/smj

from accumulated direct operating experience and reflects learning-by-doing. The coefficient λ here is a measure of the importance of direct experience in productivity improvement. This definition also enables us to isolate learning-by-doing from other sources of productivity improvement. The first term, Φ_{jt}, captures any industrywide improvements in productivity (subscript 'j' refers to industry). This may occur, for instance, because of innovations in the equipment used in the industry, or because of improved practices that become available to all firms in the industry. The last term, ν_{ijt}, captures any improvements in productivity resulting from firm-specific factors other than learning-by-doing.

Like the traditional learning curve, this approach is purely empirical and does not delve into the mechanisms of learning or even changes in firm behavior as a result of learning. Rather, it attempts to measure learning by attributing observed changes in firm performance to an observable proxy for prior experience. Though a very simple and stylized representation of the complex learning processes at play, we believe that the coefficient λ so obtained can reasonably be interpreted as the importance of prior experience X in the production process. It also offers a number of other advantages, some of which are specific to our context.

First, our study is set in the U.S. manufacturing sector, and it stands to reason that manufacturing processes would be important in determining overall firm performance. Hence, the notion of a 'production function' makes intuitive sense, and focusing on the importance of experience in the production process or on productivity improvement as a measure of learning is meaningful. Another advantage of this approach is that it controls for efficiency gains resulting from economies of scale. A traditional learning curve includes only the cumulative output, which could easily proxy for the scale of production (Argote, 1999: 16). By including current levels of physical inputs in the specification, the production function controls for the possibility that economies of scale (which is a relation between *current* output and *current* inputs) rather than learning-by-doing (which depends on *past* output) is driving improvements. As explained above, this approach also allows us to control for the possibility that improvements in manufacturing processes are a result of industrywide improvements in technology rather than direct experience. Also, under some reasonable assumptions, Equation 8 is approximately equivalent to the traditional unit cost learning curve. Finally, compared to a traditional learning curve formulation, Equation 8 involves variables that are more easily available. The main disadvantage is that these variables are usually available only at the plant level and not for individual products. Hence, the learning estimates obtained using this approach represent an average learning rate across products manufactured within a plant.[4]

DATA AND EMPIRICAL ESTIMATION

Data

The data for this study comes from two sources: Compustat and the USCB. There are two stages of analyses in this paper. First, we use *plant-level* data from the USCB to estimate the learning coefficients for each industry. We then employ these estimated *industry* learning coefficients as independent variables in regressions that use Compustat data to explore the impact of learning intensity on the heterogeneity of firm performance. These two data samples are described below.

First-stage plant-level sample (USCB data)

The first-stage sample is obtained from confidential microdata available at the USCB. Since 1972, the USCB has conducted a census of manufacturing (CM) every five years. (There were two previous censuses in 1963 and 1967.) These censuses collect detailed *plant*-level data from all U.S. manufacturing establishments with at least one employee. The data collected generally include the value of plant shipments, materials and energy inputs, employment, production hours, payments to labor, book values of physical assets, capital expenditures, inventories, and ownership (single plant firm versus part of a multiplant firm). In addition, the USCB also performs an Annual Survey of Manufactures (ASM) that collects similar data from a sample of U.S. manufacturing establishments. In particular, the annual surveys are designed to obtain an overview of the sector during the intercensal years, and hence place considerable weight on large plants and plants belonging to multiplant

[4] As we explain later, we select our sample in such a way that we reduce the possibility of very different products being manufactured in the same plant.

firms. To account for new entrants, a sample of new entrants is added to the ASM sample every year.

The USCB has collated the data from all these censuses and surveys and linked them through a longitudinal identifier to create a dataset (sometimes called the Longitudinal Research Database, or LRD), which it makes available to researchers at Census Research Data Centers, subject to access restrictions and disclosure constraints. The most important disclosure constraint is that no data that can identify or relate to a single firm or plant can be disclosed. Hence, in this paper, we do not identify statistics such as the median, minimum, or maximum for variables obtained using USCB data. For further details on this dataset, CM, or ASM, please refer to the USCB Web site.

Our sample is drawn from the LRD, which contains over 4 million plant-year observations from 1963 to 2001. Since the USCB expends more effort on larger plants and firms, the quality of data for such cases is better, and they tend to have greater continuity of observations over time. To ensure reasonable data quality, we apply some sample selection criteria, the most important of which are:

- Eliminating all plants that were established before 1973 or after 1997. Because 1973 is the first year of the annual ASM, it is not possible to reliably obtain the entry year for plants that first appear in the 1963, 1967, or 1972 censuses. In 1997, the USCB switched from the SIC to the NAICS (North American Industry Classification System). Hence, we excluded plants established after 1997 to minimize errors from industry misclassifications.[5]
- Excluding all subsequent observations for a plant if the gap between two consecutive survey years for that plant is more than two years. This is done to ensure a higher reliability of our main variable, prior cumulative output.
- Removing all plants that have a primary industry specialization ratio (the output share of the primary four-digit SIC industry in the case of a multiproduct plant) of less than 75 percent. This is done to ensure homogeneity within an industry.
- Dropping outlier plants that are in the top 0.5 percentile of capital-labor ratio or of growth in number of employees, shipments, or capital expenditure.

The resulting sample contains 182,603 plant-year observations. Summary statistics for this sample are provided in Table 1a.

Second-stage sample (Compustat data)

We use Compustat (limited to firms that have a strictly positive total asset value) to obtain data for testing the relationship between learning by doing and performance heterogeneity. This sample is obtained by aggregating firm-year level data to the industry-year level. First, for each firm-year observation, we compute Tobin's q as the ratio of market value of assets to book value of assets, and profitability as the ratio of operating profits before depreciation to total assets.[6] We then eliminate all outlying observations in the top and bottom one percent in terms of firm q or firm profitability. These data on firm performance are then aggregated to obtain the dispersion in firm q and firm profitability for each three-digit SIC industry in each year from 1973 to 2000. We also obtained other industry level variables such as industry R&D and advertising intensity from Compustat. The industry classification was based on the primary industry code. The resulting sample contains 1,523 industry-year observations. Summary statistics for this sample are included in Table 1b.

Variables

The important variables used in this study are described below. The first six pertain to the first-stage plant-level sample, and the last relates to the second-stage industry-year sample.

Output: Output for any plant for any year prior to 1996 is generally defined as the sum of the value of the plant's shipments (total plant revenues, deflated using four-digit SIC industry-year deflators available on the National Bureau of Economic Research Web site) and the difference between year-beginning and year-ending deflated work in process and deflated finished goods inventories.[7]

[5] Older plants that continued after 1997 were assumed to have retained the same four-digit SIC code they had in 1997.

[6] The definition of market value follows Kaplan and Zingales (1997).

[7] This definition is identical to that implicitly used by the USCB in its computations of plant 'value added' (see below). The

DOI: 10.1002/smj

Table 1a. Overall descriptive statistics (census or first-stage sample)[a]

Variable		Mean	S.d	Min	Max	1	2	3	4	5
1.	Value added[b]	7.40	1.86							
2.	Capital	6.62	2.12			0.77***				
3.	Labor	4.59	1.59			0.85***	0.75***			
4.	Material	7.29	2.12			0.81***	0.77***	0.79***		
5.	Prior experience	9.21	2.24			0.83***	0.83***	0.80***	0.85***	
1b Overall descriptive statistics (Compustat or second-stage sample)										
1.	Industry R&D intensity	0.02	0.02	0.00	0.12					
2.	Industry advertising intensity	0.02	0.02	0.00	0.08	−0.08***				
3.	Number of firms	44.15	51.31	10	453	0.72***	−0.09***			
4.	Industry sales ($ billion)	47.59	140.40	0.44	1,498	0.12***	−0.10***	0.26***		
5.	Industry q range (10^{th}–90^{th} pctile)	1.65	1.36	0.093	10.86	0.47***	0.16***	0.41***	0.03	
6.	Industry profitability range (10^{th}–90^{th})	0.28	0.16	0.03	1.59	0.39***	0.07***	0.38***	−0.01	0.63**

[a] There are *two* separate samples. Table 1a refers to the first-stage plant-level sample (n = 182,603 of which 170,666 are in three-digit SIC code industries that have at least 50 plants) for which we are not able to present the minimum and maximum due to disclosure restrictions. Table 1b refers to the second-stage sample based on Compustat data (n=1,523)
[b] Variables 1–5 in Table 1a are logarithms of their original values. Please refer text for precise definition of variables.
* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Due to the unavailability of inventory data, output for the years including and after 1996 is simply defined as the deflated shipments.

Value added: Value added is defined as the difference between real output and real materials (described below).

Labor: We define quantity of labor to be the labor hours expended in production worker equivalents. Labor hours for any plant are computed by dividing the total wage bill for the establishment by the average hourly wage for production workers in that establishment.

Materials: Real materials are defined as the sum of deflated cost of material purchases, external contract work, fuel, and electricity.

Capital stock and capital investment: We use the perpetual inventory approach to compute real capital stock. We compute separate stocks for buildings (or structures) and machinery. Real capital stock (k_{it}) in any given year, say for machinery, is computed as $k_{it} = (1 - d)k_{it-1} + I_{it-1} + R_{it}$, where d is an industry-year specific depreciation rate for machinery, I is the capital investment in machinery (deflated by an industry-year specific investment deflator for the year $t - 1$) and R is the capitalized value of capital equipment rentals. If an establishment is not observed every year, following Olley and Pakes (1996) we impute gross investment linearly (i.e., $I_{it} = 0.5 \times (I_{it} + I_{it-n}) \times (n - 1)$, where I_{it} is the imputed investment for period t and n is the gap between the two survey years).

Prior cumulative output: This is used as a proxy for accumulated operating experience. Prior cumulative output is defined as the sum of real output through the end of the previous period, that is, $X_{it} = \mathrm{sum}(o_{i1},\ o_{i2} \ldots o_{it-1}) = X_{it-1} + o_{it-1}$, where o is real output. If an establishment is not observed every year, we impute output linearly (i.e, $\mathbf{O}_{it} = 0.5 \times (o_{it} + o_{it-n}) \times (n - 1)$, where $\mathbf{O}_{it}$ is the imputed output for period t and n is the gap between the survey years).[8]

USCB uses slightly different definitions in some industries due to differences in the nature of the manufacturing processes. We follow the USCB's definitions in all these cases. A detailed description is available on request.

[8] This measure of experience does not incorporate 'organizational forgetting' and hence, does not differentiate between a small, old firm and a large, young firm. As rough robustness checks, we estimated the learning coefficients (i) using the cumulative output through t-2 as a measure of experience, and (ii) including plant age as another variable in Equation 8. The learning coefficients so estimated were highly correlated with our baseline estimates. We also used a nonlinear specification incorporating organizational forgetting as a rough robustness check and found those estimates to be highly correlated with our baseline estimates.

DOI: 10.1002/smj

Heterogeneity of firm performance: We use firm q and firm profitability as measures of firm performance. Unlike most studies in the literature, we use direct measures of performance heterogeneity, specifically the cross-sectional dispersion of firm performance. We use three measures of cross-sectional dispersion of firm q and firm profitability. As the baseline measure, we take the difference between the 90th percentile and 10th percentile (of firm q or profitability) in an industry during a given year. The advantage of using this measure as opposed to, say, variance, is that it is an ordinal measure and hence much less affected by the presence of outliers. As robustness checks, we use the interquartile range (or the difference between the 75th and 25th percentiles) and the standard deviation of firm profitability (or q) as other measures of heterogeneity.[9]

Empirical estimation

In the first part of our study, we use the census data to estimate the learning rates for each three-digit SIC industry and to characterize the heterogeneity in industry learning rates. In the second part, which addresses the link between industry learning intensity and the heterogeneity of firm performance, we use the estimated industry-by-industry learning coefficients as explanatory variables in regressions with the range of firm performance as the dependent variable.

Measuring industry learning intensity: To proceed with empirical estimation of the importance of learning, we use ordinary least squares (OLS) to estimate the logarithmic version of Equation 8:

$$y_{ijt} = a_{jt} + \alpha_j.k_{ijt} + \beta_j.l_{ijt} + \lambda_j.x_{ijt} + \varepsilon_{ijt} \tag{9}$$

where y is log (value added); k, l, and x are log(L), log(K) and log(X) respectively; a is log(φ); and ε includes log(ν).

The coefficient of interest is λ_j, the learning intensity of industry 'j'.[10] We estimate Equation 9 for *each* three-digit SIC industry that has more than 50 plants. Estimating the production function industry-by-industry ensures that we are excluding the possibility that differences in returns to scale are being spuriously captured as learning. The terms a_{jt} in Equation 9 are coefficients on industry-year dummies, which capture all intertemporal movements in the average industry productivity, including any industrywide technology improvements. Hence, the econometric identification of the coefficients comes solely from cross-sectional deviations from the industry-year averages and not from changes in mean industry productivity over time.[11]

We conclude this subsection with a brief comment on the link between the Jovanovic and Nyarko (1995) model described earlier, and the empirical estimation approach explained above. By estimating the production function (Equation 9), we are approximating the logarithmic version of Equation 5. This is also equivalent to approximating the logarithmic version of Equation 3, since the maximum efficiency term in Equation 5 is an industry characteristic, and hence will drop out when industry-year fixed effects are included.

[9] Note that the relation between the heterogeneity of these performance measures and that of efficiency (the variable of interest in the Jovanovic and Nyarko (1995) model discussed earlier) is not necessarily monotonic. A greater dispersion of efficiency (due to learning) does not always imply a greater dispersion of profits and market value. More efficient firms will generally have higher profits and market value than inefficient firms in the same industry. (Thus, this implies that an increase in dispersion of efficiency will increase the dispersion of profit, all other things being equal). However, the presence of learning may decrease the industry price, and force some of the inefficient firms to exit (thereby decreasing dispersion). To achieve a monotonic relationship, it needs to be the case that the profit function is strictly increasing (and bounded) in efficiency, and that the efficiency gains from learning are higher for more efficient firms.

[10] Since we treat the learning environment to be an industry characteristic, we estimate only one learning coefficient per industry. However, learning rates may change with time. As a robustness check, we estimated separate learning rates for the periods 1973–1984 and 1985–2000 (roughly equal subsamples). The Spearman rank correlation between these two sets of coefficients was 0.54 and between these coefficients and our baseline estimates 0.80 and 0.87, respectively, all statistically significant at or below the 0.01 percent level. Further, we estimated learning rates separately for plants established early versus late in our sample (plants established during 1973–1984 versus 1985–2001). These learning rates were highly correlated with the baseline rates. Similarly, the inclusion of controls for age and entry-year fixed effects resulted in learning rates highly correlated with the baseline rates.

[11] One potential concern could be the bias in OLS estimates arising from the endogeneity of input choices, and survival bias. In order to address these concerns, we developed extensions of the Olley and Pakes (1996) and Ackerberg, Caves, and Frazer (2006) methods of production function estimation to estimate learning rates. Those estimates were highly correlated with the OLS estimates. More importantly, our primary results regarding firm performance heterogeneity remain robust to these alternative estimates (results available on request).

DOI: 10.1002/smj

Unlike the pattern described by Equation 3 or Equation 5, which rise asymptotically to a maximum level of efficiency, the production function estimates a linear function (in logarithms). While we could estimate production functions with additional nonlinear terms (refer to the robustness checks section for a translog production function), we use the simple linear version for several reasons. First, it is in line with most prior empirical studies of learning curves. A single learning rate for each industry also makes it easier to understand differences across industries and rank industries based on learning rates. Moreover, the model outlined above is only a stylized version of learning processes, and a linear function is the natural first approach to approximating the processes.

Finally, the empirically *estimated* learning rate will be affected by the distribution of experience within the industry. For any given set of parameter values, industries with more mature firms will tend to exhibit lower *estimated* learning rates than industries with younger firms. We performed some indicative simulation analysis of the link between the model and the estimated learning rate. Specifically, we estimated the learning rate by fitting a linear function to Equation 3 under different *assumed* parameter values and experience distributions. These simulations (available on request) suggest that for any given set of parameter values, an industry composed entirely of mature firms would demonstrate a learning rate roughly one-third lower than an industry with firms distributed evenly across experience cohorts, and an industry with predominantly young firms would demonstrate a learning rate about 10 percent higher. The simulations also suggest that such variation is small compared with the impact of changes in the complexity parameters. Moreover, our data do not show significant interindustry differences in the age distribution. The mean plant age in an industry varies from 3.08 years to 9.53 years, and the mean within-industry variance in age is comparable to the variance in the mean age across industries (4.04 years versus. 5.19 years). The distribution of experience exhibits a similar pattern. Hence, interindustry differences in plant age or experience are unlikely to have much effect on the estimated rates of learning.

Interindustry heterogeneity in learning: It is difficult to obtain good empirical measures of process complexity for such a large sample. Hence, we rely on rather simple proxies that are easily available. Industries with greater capital-intensity have been associated with greater process complexity (Lieberman, 1984). Also, it is reasonable to believe that R&D-intensive industries have more complex processes and involve a higher degree of knowledge tacitness. Similarly, industries with high advertising intensity are likely to be differentiated, thus reducing the amount of learning that firms can achieve from their competitors. Finally, industry wages may reflect the underlying skill requirements, and higher wages may proxy complexity.

To formally test Hypothesis 1, we adopt the following regression model, which includes these proxies interacted with x_{ijt}, the cumulative output measure:

$$\begin{aligned} y_{ijt} = a_{jt} + \alpha.k_{ijt} + \beta.l_{ijt} + \lambda.x_{ijt} + \lambda_1.C_{jt.}x_{ijt} \\ + \lambda_2.W_{jt.}x_{ijt} + \lambda_3.R_{jt.}x_{ijt} + \lambda_4.A_{jt.}x_{ijt} \\ + \varepsilon_{ijt} \end{aligned} \tag{10}$$

where C is industry capital intensity (capital stock to employment ratio); W is industry wages; R is industry R&D intensity (R&D expenditure divided by sales); and A is industry advertising intensity (advertising expenditure divided by sales).

As in the industry-specific learning regressions, the unit of analysis is plant year, and we allow for industry-year dummies a_{jt}. For this analysis, industry R&D and advertising data are obtained from Compustat. We use OLS to estimate Equation 10, with plant fixed effects and instrumental variables specifications as robustness checks.

Industry learning intensity and heterogeneity of firm performance. In order to examine how industry learning is related to heterogeneity of firm performance (Hypothesis 2), we use regressions of the following form:

$$\begin{aligned} \pi_{jt} = a_t + b.\underline{\underline{\lambda}}_j + c_1.R_{jt} + c_2.A_{jt} + c_3.C_{jt} \\ + c_4.S_{jt} + c_5.N_{jt} + c_6.P_{jt} + \varepsilon_{jt} \end{aligned} \tag{11}$$

where π_{jt} is 90^{th} to 10^{th} percentile range of firm performance, either firm q or firm profitability, in industry j during year t; $\underline{\underline{\lambda}}_j$ is the estimated industry learning intensity from Equation 9; R is industry R&D intensity (R&D expenditure/sales); A is

DOI: 10.1002/smj

industry advertising intensity (advertising expenditure/sales); C is industry capital intensity (total assets/sales); P is average industry profitability (operating profits/total assets); N is the number of firms in an industry; and S is industry size measured as total industry sales.

Note that the level of analysis here is the *industry year*. For this analysis, we rely on data from Compustat; the only variable in Equation 11 that comes from outside Compustat is the estimated learning intensity.

A brief note on the choice of control variables is in order. Intuitively, our earlier arguments on the interindustry heterogeneity in learning rates also apply to any factor that increases complexity. R&D, advertising, and capital intensity can logically be classified as such factors and would tend to increase the heterogeneity of firm performance. For instance, industries with high R&D or high advertising intensity may be quite differentiated and hence, performance more dispersed. Furthermore, these are sunk costs, which increase the incentives for firms to stay in the market once they have incurred those costs (Gschwandtner and Lambson, 2006), thereby increasing interfirm heterogeneity. Average industry profitability may reflect inherent risk and hence may be associated with a higher variance of returns. Finally, we add the number of firms and industry size as factors that may potentially increase measured heterogeneity.

RESULTS

Interindustry heterogeneity in learning-by-doing

Table 2 presents the results of estimating Equation 9 for the *pooled* sample. Model 1 is a simple Cobb-Douglas production function, excluding the prior experience term that captures learning-by-doing. Model 2 expands on Model 1 by adding the prior experience term. The coefficient on prior cumulative output is 0.26, which implies a 19.7 percent gain in productivity for each doubling of cumulative output.[12] This model, however, does not control for the possibility that the rate of technological improvement varies across industries. For instance, firms in an industry with significant technological advances may show productivity improvements even without learning-by-doing. A robust approach to address this is to include a dummy variable for each industry-year combination, which will control for all intertemporal changes (including technological improvements) in the average industry productivity. While our baseline definition of industry is at the three-digit SIC level, the size of our pooled sample permits us to follow a far more conservative approach and use four-digit SIC industry-year fixed effects. Model 3 includes 9,967 separate four-digit SIC industry-year dummies, which control for all productivity improvements in each four-digit SIC industry (and consequently, in each three-digit SIC industry). The estimated learning coefficient falls to 0.23 when these controls are added.

We then estimated Equation 9 using OLS for each of the 117 three-digit SIC industries that has more than 50 plants. Models 4-1 to 4–117 allow each industry to have its own coefficients on capital, labor, and prior cumulative output. They also allow for year dummies within each three-digit SIC industry and hence, control for three-digit SIC industry-wide productivity improvements. Given space constraints, we present only the coefficients on cumulative output from these models in Appendix A (Column 3). Figure 3 presents a histogram of the coefficients on cumulative output. As expected, there is a significant variation in learning intensities across industries, ranging from just above zero to almost 0.60 with an average of 0.22 (almost identical to the estimate for the pooled sample in Model 3).

We now try to characterize the heterogeneity in learning intensity. From Appendix A, we can see that the top six industries based on the OLS learning coefficient are SIC 357 (computers), 283 (pharmaceuticals), 291 (petroleum refining), 386 (photographic equipment and supply), 287 (agricultural chemicals), and 289 (miscellaneous chemicals). The lowest in the list are SIC 317 (leather goods), 322 (glass products), 262 (paper mills), 228 (yarn and thread mills), and 311 (leather tanning). This list suggests a positive association between complex, knowledge-intensive and capital-intensive settings, and the rate of learning-by-doing.

To test this more formally, we estimate Equation 10 using OLS. Again, our pooled sample permits us to adopt a more conservative approach and use a finer four-digit SIC industry definition. The variables of interest are the interaction terms between

[12] This is computed as $2^{0.26}$.

DOI: 10.1002/smj

Table 2. Pooled learning coefficients

	Variable	Model 1 (OLS)	Model 2 (OLS)	Model 3 (OLS)	Models 4-1 to 4-117 (OLS)
1.	Capital	0.271*** (0.003)	0.12*** (0.003)	0.07*** (0.003)	−0.12 to 0.29 (Details available on request)
2.	Labor	0.710*** (0.004)	0.59*** (0.004)	0.66*** (0.004)	0.34 to 0.98 (Details available on request)
3.	Prior experience		0.26*** (0.004)	0.23*** (0.004)	0.00 to 0.60 (Details available in Appendix A)
4.	Fixed effects	Year	Year	Four-digit SIC industry year	Three-digit SIC industry year
N		213,256	170,666	170,666	83 to 7,244 (Details available on request)
R^2		0.80	0.81	0.85	
Adjusted R^2		0.80	0.81	0.85	

[a] The unit of analysis is *plant year*. Value added is the dependent variable. Coefficients on dummies not presented.
[b] Models 4-1 to 4-117 are 117 separate OLS estimations along the lines of Model 2, one for each three-digit SIC code industry.
* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Robust standard errors in parentheses.

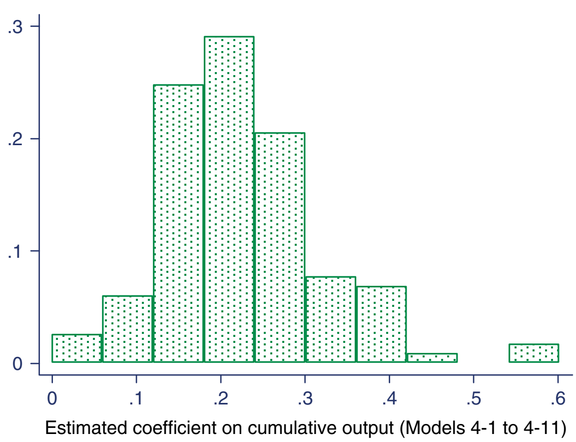


Figure 3. Interindustry heterogeneity in learning-by-doing. This figure is available in color online at www.interscience.wiley.com/journal/smj

prior experience and industry factors. Models 5 and 6 in Table 3 use a larger sample for which we have complete data on industry wages and the capital-labor ratio, omitting the industry R&D and advertising intensity terms. Model 5 includes only year indicators while Model 6 includes a full set of industry-year dummies. In both cases, the learning coefficient is significantly higher in industries with greater capital intensity. The interaction effect of industry wages on prior experience becomes insignificant once industry-year effects are controlled for.

Model 7 estimates Equation 10 with a smaller sample for which we have complete industry R&D and advertising data from Compustat. The coefficient on prior experience is significantly higher in industries with higher capital-labor ratio and greater R&D and advertising intensity. Models 5–7 assume that the coefficients on capital and labor are the same across industries; Model 8 repeats the tests in Model 7, allowing the coefficients on capital and labor to vary across two-digit SIC industries. The results in Model 8 are not substantially different from Model 7. Finally, we estimate Models 9 and 10 to check if these results are robust to the inclusion of plant fixed effects. Model 9 does not include any of the direct terms while Model 10 includes them. While the effects decrease considerably in magnitude (as expected) in Model 9, the direction and statistical significance of the results persist. In Model 10,

DOI: 10.1002/smj

Table 3. Interindustry heterogeneity in learning-by-doing[a]

		Model 5 (OLS)	Model 6 (OLS)	Model 7 (OLS)	Model 8[b] (OLS)	Model 9[b] (FE)	Model 10[b] (FE)
1.	Capital	0.064***	0.074***	0.077***	By two-digit SIC code	By two-digit SIC code	By two-digit SIC code
		(0.003)	(0.003)	(0.005)			
2.	Labor	0.647***	0.665***	0.667***	By two-digit SIC code	By two-digit SIC code	By two-digit SIC code
		(0.004)	(0.004)	(0.007)			
3.	Prior experience	0.250***	0.225***	0.198***	0.191***	0.066***	0.022*
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
4.	Experience* industry wages[c]	−1.14***	−0.526	0.083	0.069	0.742***	0.036
		(0.31)	(0.37)	(0.52)	(0.57)	(0.14)	(0.475)
5.	Experience* industry capital labor ratio[c]	0.428***	0.323***	0.226***	0.266***	0.049***	0.424***
		(0.05)	(0.32)	(0.06)	(0.07)	(0.01)	(0.04)
6.	Experience* industry advertising intensity			0.869***	0.946***	0.156***	0.553***
				(0.14)	(0.16)	(0.03)	(0.14)
7.	Experience* industry R&D intensity			0.446***	0.503***	0.037***	1.35***
				(0.10)	(0.11)	(0.01)	(0.08)
8.	Industry wages	0.044***					0.004
		(0.003)					(0.01)
9.	Industry capital labor ratio[c]	−0.036***					−0.446***
		(0.004)					(0.04)
10.	Industry R&D intensity[d]						−0.149***
							(0.01)
11.	Industry advertising intensity[d]						−0.047***
							(0.02)
12.	Fixed effects	Year	Industry year	Industry year	Industry year	Plant and year	Plant and year
N		182,603	182,603	71,824	71,824	71,824	71,824
R^2		0.82	0.83	0.83	0.83	0.91	0.93

[a] The unit of analysis is plant year. Log value added is the dependent variable. Coefficients on dummies not presented. With industry-year effects, direct terms are not included because the dummies control for all changes at the industry-year level (e.g., industry R&D intensity).
[b] Models 8 to 10 allow for *industry-specific* (by two-digit SIC code) capital and labor coefficients. Given the large number of capital and labor coefficients, we do not present them here.
[c] Coefficients and standard errors divided by 1,000 for presentation purposes.
[d] Coefficients and standard errors divided by 100 for presentation purposes.
* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Robust standard errors appear in parentheses.

the significance of the interaction terms increase considerably and the direct terms are negative. This suggests that in industries with high capital, R&D, or advertising intensity, plant productivity is initially low but rises steeply with experience. Finally, we used an instrumental variables specification with once-lagged variables as instruments. The economic substance and statistical significance of these results (available on request) were very similar to those in Models 7 and 8.

To summarize, we find that the learning rate increases with industry capital intensity, R&D intensity, and advertising intensity, thus confirming Hypothesis 1.

Industry learning intensity and firm performance heterogeneity

We now examine the important question of how these variations in industry learning intensity are related to firm performance heterogeneity. Figure 4 gives an idea of the potential linkages. It presents the distributions of firm profitability and firm q (relative to the industry-year average) for 'high learning' (i.e., industries with learning rates above the median learning rate) and 'low learning' industries. Figure 4 shows that both measures of firm performance have greater dispersion in high learning industries. To test this formally, we use the industry estimated learning coefficients (shown in Appendix A) as independent variables in Equation 11, with the range of firm performance within an industry as the dependent variable. Recall that the unit of analysis in Equation 11 is the industry year.

Table 4 presents the results of estimating Equation 11 for our two measures of performance heterogeneity. Since all the variables except the learning coefficient have very skewed distributions, we use their logarithms rather than the original values. Model 11 uses the range of firm profitability as the dependent variable while Model 12 uses the range of firm q. The industry learning intensity coefficient is 0.926 in Model 11. This implies that the difference between the 'best performers' (top 10%) and the 'worst performers' (bottom 10%), as measured by relative profitability, is considerably greater in industries with high learning. Similar results hold when we use firm q as a measure of firm performance. In both cases the coefficient on industry learning is positive and statistically significant at the one percent level. These results provide strong support for Hypothesis 2, and demonstrate that the heterogeneity of firm performance increases with industry learning rate.

Turning to the other coefficients in the regressions, sunk costs (particularly R&D and advertising, and to an extent, capital intensity) are also positively linked to increased dispersion of firm performance. This is broadly in line with Gschwandtner and Lambson (2006), who found that sunk costs tend to increase profit variability in an industry, and with theoretical industry models in the economics literature such as Hopenhayn (1992) that predict increased productivity dispersion due to higher sunk costs.[13] The mean industry profitability appears to be associated with increased heterogeneity of firm value (which is consistent with a higher risk-higher return story). Counterintuitively, mean industry profit is negatively associated with the dispersion of firm profit, but this could be an accounting artifact due to the inclusion of depreciation within the measured profit rate. Another counterintuitive result is that large industries (by industry sales) tend to have lower heterogeneity. We do not have a good explanation for this except, perhaps, that they may be mature industries. Industries with many firms, in line with our intuition, show a wider dispersion.

Robustness checks

We performed a series of tests to confirm that we are most likely measuring the effect of learning-by-doing and that our subsequent results on the heterogeneity of firm performance are robust to alternative specifications. Briefly, the tests show that factors such as survivor bias, sample selection, R&D investments, measurement errors in capital, choice of production function form, and industry life cycle effects are not driving the observed heterogeneity in learning rates. Details are provided in Appendix B and Table 5.

We tested the robustness of our results on the connection between learning and firm performance by running the same type of regressions as in Table 4, but with different measures of performance heterogeneity, levels of aggregation,

[13] In his working paper, Rivkin (1998) also finds that the dispersion of firm profit rates is higher in industries with opportunities for R&D and product differentiation.

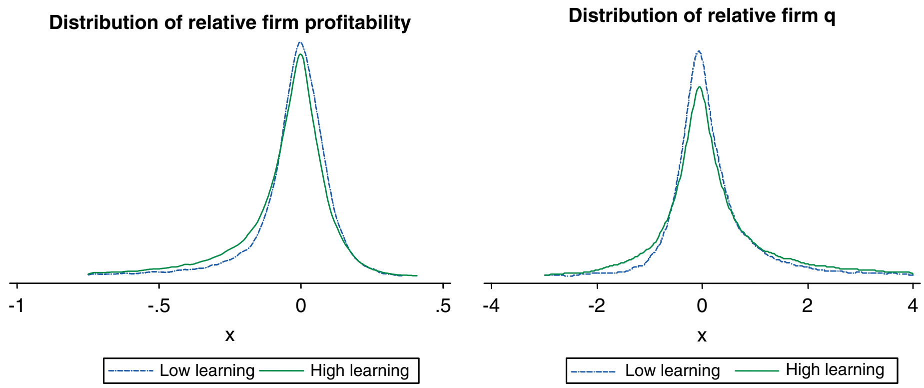


Figure 4. Industry learning environments and firm performance heterogeneity. This figure is available in color online at www.interscience.wiley.com/journal/smj

Table 4. Learning-by-doing and heterogeneity in firm performance[a]

Dependent variable		Model 11	Model 12
		Profit dispersion	*q* dispersion
1.	Industry learning intensity	0.926*** (0.10)	1.41*** (0.14)
2.	Industry R&D intensity	0.050*** (0.01)	0.102*** (0.01)
3.	Industry advertising intensity	0.040*** (0.01)	0.070*** (0.01)
4.	Industry profitability (mean)	−0.131*** (0.04)	0.207*** (0.05)
5.	Industry capital intensity	−0.110** (0.05)	0.113* (0.06)
6.	Industry sales	−0.117*** (0.01)	−0.117*** (0.01)
7.	Industry number of firms	0.251*** (0.02)	0.268*** (0.02)
8.	Fixed effects	Year	Year
N		1,523	1,523
R^2		0.38	0.55
Adjusted R^2		0.37	0.54

[a] The unit of analysis in all these regressions is three-digit SIC industry year. All variables except the industry learning intensity are logarithms of their original values. The dependent variable is the (logarithm of) difference between 90th and 10th percentiles of either firm profitability or firm q. Coefficients on dummies not presented.
* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Robust standard errors appear in parentheses.

choice of time periods, assumptions about error correlation structures, and others. Table 5 presents the coefficients on the learning estimates from those regressions. (Each line in Table 5 is comparable to line 1 from Table 4; details available on request). In all of these regressions, learning shows a significant positive association with the heterogeneity of firm performance.

DISCUSSION AND CONCLUSION

It is widely accepted and documented in the strategy literature that the industry environment affects competitive heterogeneity. In addition, many studies have shown that the rate of organizational learning varies greatly across firms and industries. Nevertheless, given data limitations the connections between learning intensity and the industry competitive environment have not been systematically explored. Our paper provides quantitative evidence that supports the case for treating learning intensity as a fundamental characteristic of the industry, much as R&D and advertising. Furthermore, our paper facilitates empirical implementations of such a concept by providing reasonably comparable estimates of learning intensity for a wide range of industries, encompassing most of the U.S. manufacturing sector.

To our knowledge, this study is the first to provide a quantitative comparison of learning rates across such a broad set of industries. The range of learning slopes (slope computed as $2^{-\lambda}$, where λ is the estimated learning coefficient) varies from close to 100 percent (or no learning) in SIC 262 (paper mills), to a maximum of 68 percent in SIC 283 (drugs) and 66 percent in SIC 357 (computers).[14] This is comparable to the range of estimates in prior studies. For instance, the survey by Dutton and Thomas (1984) found that the median learning slope in about 22 industry-specific studies was about 80 percent and the range was from 55 percent to 110 percent.

Our study goes beyond simply establishing the heterogeneity in learning rates to identify some broad patterns in these rates, as indicated in Table 3. Even within the limited interpretation permitted by our crude proxies, the results are consistent with the argument that learning from own experience may be more important in environments where complexity is high. Knowledge transfers between firms, and perhaps even within firms, is naturally harder in such environments, and firms may have to rely more on their experience. Though intuitive, this study is the first attempt to quantify these patterns in a systematic way across a broad sample of industries.

The second contribution of our paper is to demonstrate a robust association between the industry learning intensity and the cross-sectional heterogeneity of firm performance. The results in Table 4 show that firm performance is more heterogeneous in high learning industries. More importantly, the economic significance of this effect seems to be large. Based on the coefficients in Model 11, an increase of one standard deviation in the learning coefficient (0.097) is associated with a 31 percent ($= 0.926 \times 0.097/0.2891$) increase in the profitability range.[15] These are comparable to or even higher than the effect of an increase in R&D or advertising intensity. For instance, using coefficients from Model 11, an increase in R&D intensity by one standard deviation from the mean (i.e., from 0.0217 to 0.0443) is associated with a 12 percent ($= 0.050 \times [\ln(0.0443) - \ln(0.0217)]/0.2891$) increase in profitability range.

Thus, we have shown that the cross-sectional variation in firm performance increases with the learning intensity of the industry. This is consistent with the predictions of the Jovanovic and Nyarko (1995) model presented in the early part of this paper. More broadly, the evidence points to a story more like the uncertain learning-by-doing in Levitt and March (1988) and Levinthal and March (1993) rather than the sure shot learning curve often assumed in the economics literature. If uncertain learning creates winners that grow to be bigger than others, the *size-weighted* average firm performance should be significantly higher in industries with high learning. Indeed, the asset-weighted industry profit ratio is about 14 percent in high learning industries compared to 11 percent in low learning industries. The asset-weighted industry average q is 1.29 in high learning industries, versus. 1.08 in low learning industries.

Although we have focused on how learning intensity affects the variance of firm profits and q, Figure 4 suggests that learning intensity may also impact the skewness of these distributions. Regressions similar to those in Table 4 with the top and bottom deciles of the firm profit distribution as the dependent variable show that the increased dispersion in firm profits is almost entirely due to a downward shift in the bottom10th percentile of profits rather than an upward shift in the 90th percentile. In contrast, the heterogeneity of firm q appears to be mostly due to an upward shift in the 90th percentile rather than a decrease in the 10th percentile.[16] These results imply that firms in industries with high learning intensity may show losses during early years of operation, though

[14] A learning slope of x percent implies that a doubling of cumulative output leads to a (100-x) percent increase in productivity.

[15] Similar calculations using the Olley and Pakes (1996) and Ackerberg *et al*. (2006) methods yield estimates of 24 percent and 29 percent (results available on request).

[16] The unweighted industry average profit is 9.58 percent for low learning industries and 7.08 percent for high learning industries. The corresponding unweighted industry average *q*s are 1.01 and 1.42.

Table 5. Learning-by-doing and heterogeneity in firm performance: robustness checks[a]

Dependent variable		Profit dispersion	q dispersion
1.	75th to 25th percentile as dependent variable	0.717*** (0.10)	1.34*** (0.13)
2.	Standard deviation as dependent variable	0.170*** (0.02)	1.32*** (0.14)
3.	Learning ranks instead of coefficients[b]	0.396*** (0.06)	0.649*** (0.08)
4.	Two learning categories - high versus low instead of coefficients (based on median learning rate)	0.123*** (0.02)	0.187*** (0.03)
5.	Including only very focused firms[c]	1.01*** (0.20)	1.05*** (0.24)
6.	Four-digit SIC as industry definition	0.452*** (0.11)	0.951*** (0.16)
7.	Excluding industries ending with '9' and including # of four-digit SIC codes as control	0.826*** (0.11)	1.25*** (0.15)
8.	Without taking logarithms	0.264*** (0.04)	2.00*** (0.32)
9.	Dependent variable not logged	0.351*** (0.04)	2.87*** (0.31)
10.	Excluding industry years with <25 firms	0.888*** (0.13)	1.26*** (0.16)
11.	Clustering standard errors at three-digit SIC level[d]	0.926*** (0.21)	1.41*** (0.27)
12.	Excluding the two highest and two lowest learning industries	0.726*** (0.16)	1.28*** (0.21)
13.	Including two-digit SIC year fixed effects	0.974*** (0.14)	1.39*** (0.20)
14.	Time period 1973–1984[e]	0.640*** (0.17)	2.40*** (0.23)
15.	Time period 1985–2000[e]	0.613*** (0.15)	1.50*** (0.21)
16.	Learning coefficients with R&D controls[f]	0.349*** (0.06)	0.633*** (0.08)

[a] This table provides the results of robustness checks that use the same type of regressions as in Table 4, but with different measures of performance heterogeneity, level of aggregation, choice of time periods, etc. Each line in this table is comparable to line 1 from Table 4. Only the coefficients and standard errors on industry learning intensity are presented. Coefficients on other variables are available on request.
[b] The coefficients and standard errors have been multiplied by 100 for presentation purposes.
[c] This regression includes only firms whose largest segment (by sales) accounts for at least 95% of the total sales. Due to nonavailability of data before 1984, this covers the period 1984 to 2000.
[d] The standard errors are computed allowing for arbitrary autocorrelation of errors within a three-digit SIC industry.
[e] These use learning coefficients estimated for the relevant time period as an independent variable.
[f] Learning coefficients used in this regression are estimated after controlling for *firm specific* R&D expenditure (the sample changes because not all firms report R&D).

expected future returns (as reflected by Tobin's q) are high.

As with all empirical studies, our analysis comes with a number of limitations. We adopt a highly aggregated view of learning by focusing on learning intensity at the industry level. Clearly, there is likely to be considerable heterogeneity in products and learning rates within industries, perhaps even greater than the interindustry variations. Moreover, our study does not shed any light on the mechanisms of learning; for example, factors within organizations such as training and engineering activities (Adler and Clark, 1991), and structures and routines (Nelson and Winter, 1982) that may affect learning. Furthermore, it is to be expected that the meaning and context of organizational learning vary significantly across (and within) industries. Since this paper follows a

DOI: 10.1002/smj

purely empirical approach and infers the importance of learning by examining the coefficient on cumulative output, the findings from this study are necessarily a very simplified and stylized representation of the learning environment. Furthermore, learning-by-doing is only one form of organizational learning (Levitt and March 1988; Malerba 1992). There are many other forms of learning, such as learning from others, which are not examined in this study. Finally, there are measurement issues that commonly afflict studies of productivity estimation. Notwithstanding these limitations, we believe that this aggregate approach provides a 'big picture' view of the heterogeneity in industry learning-environments that complements detailed microlevel studies of learning.

Although not a limitation per se, the interpretation of the learning coefficient deserves some discussion. As we measure it, the learning coefficient does not reveal two aspects of learning that have been considered in prior studies. First, it does not tell anything about spillovers of learning across firms. Firms can apply experience gained from one product to cost reduction or quality improvements in other (perhaps, similar) products (Argote, 1999; Benkard, 2000; Irwin and Klenow, 1994) and it is reasonable to expect that learning 'spills over' from one firm to another. In our approach, all improvements resulting from such industrywide learning spillovers are captured by the industry-year dummy variables. As an extreme, an industry where some firms learn considerably from experience but *all* knowledge so generated is transferred to other firms *immediately* (thereby leaving the relative performances unchanged) would be measured as having a zero rate of learning. However, in such a case, learning rates would not affect firm performance and a 'zero' learning coefficient would not entirely be meaningless. The second issue is organizational forgetting. It has been established that the knowledge accumulated through learning depreciates rapidly (Argote, 1999; Benkard, 2000). The use of a single learning coefficient clearly masks underlying differences in the rate of depreciation. A low learning coefficient could mean either a low learning rate combined with a low rate of forgetting, or a high learning rate combined with rapid depreciation. Unfortunately, our data do not permit us to reliably estimate the rates of forgetting. Even so, the learning coefficient can still be meaningfully interpreted as 'net rate of learning' or as the 'net importance' of experience in the production process.

The present study establishes basic relationships between industry learning and firm performance, but many extensions are possible. One would be to examine the mechanisms that explain the link between learning intensity and the heterogeneity of firm performance. We have provided some potential theoretical reasons for this association; however, their importance must be sorted out. More broadly, the learning estimates from this study can be used to analyze how variations in industry learning rates affect firm behavior. Strategic choices that may be affected by learning include incentive structures, governance structures, investments in innovation, capital and technology, and perhaps even organizational structures and processes. For instance, a bigger role for learning from experience may require an incentive structure oriented toward long-term performance goals rather than short-term ones. Another interesting line of inquiry would be to generalize our findings to examine how variations in the knowledge acquisition processes (e.g., through own learning versus interfirm spillovers versus intrafirm spillovers, etc.) across industries affect the observed heterogeneity.

Heretofore, researchers have been constrained by the nonavailability of learning rate data needed to address such questions across a wide range of industries. This study provides industry-specific estimates that can be used to further explore the role of learning by doing, particularly in broad, interindustry contexts. We invite others to build upon this work.

ACKNOWLEDGEMENTS

We thank an anonymous referee for comments that helped us to significantly improve our study. We are grateful to Becky Acosta at the CCRDC for her help with the project proposal and prompt processing of our disclosure requests. This paper has benefited immensely from comments by Emin Dinlersoz, Susan Cohen, and participants at the CRES-Gort and ACAC conferences, the 2006 AOM annual meeting, and the TOM area seminar at the Harvard Business School. Natarajan Balasubramanian thanks Jagadeesh Sivadasan for numerous discussions, Dan Ackerberg, Roberto Alvarez, Michael Darby, Hugo Hopenhayn, Raphael Lam, Jeongsik Lee, Gabriel Natividad, and Mariko

Sakakibara for their constructive comments and Hyowook Chiang, John Haltiwanger, and Ron Jarmin for kindly sharing various data. The CCRDC generously provided support to Natarajan Balasubramanian and covered the data access fees.

REFERENCES

Ackerberg D, Caves K, Frazer G. 2006. Structural estimation of production functions: an application to the timing of input choices. Working Paper, UCLA, http://www.econ.ucla.edu/ackerber/ACF20withtables.pdf (accessed 11 September 2009).

Adler PS, Clark KB. 1991. Behind the learning curve: a sketch of the learning process. *Management Science* **37**(3): 267–281.

Argote L. 1999. *Organizational Learning: Creating, Retaining and Transferring Knowledge*. Kluwer Academic: New York.

Bahk BH, Gort M. 1993. Decomposing learning-by-doing in new plants. *Journal of Political Economy* **101**(4): 561–583.

Benkard CL. 2000. Learning and forgetting: the dynamics of aircraft production. *American Economic Review* **90**(4): 1034–1054.

Dutton JM, Thomas A. 1984. Treating progress functions as a managerial opportunity. *Academy of Management Review* **9**(2): 235–247.

Gruber H. 1994. The yield factor and the learning curve in semiconductor production. *Applied Economics* **26**(8): 837–843.

Gruber H. 2000. The evolution of market structure in semiconductors: the role of product standards. *Research Policy* **29**(6): 725–740.

Gschwandtner A, Lambson VE. 2006. Sunk costs, profit variability and turnover. *Economic Inquiry* **44**(2): 367–373.

Hayes RH, Clark KB. 1986. Why some factories are more productive than others. *Harvard Business Review* **64**(5): 66–73.

Hopenhayn H. 1992. Entry, exit and firm dynamics in the long run. *Econometrica* **60**(5): 1127–1150.

Irwin DA, Klenow PJ. 1994. Learning-by-doing spillovers in the semiconductor industry. *Journal of Political Economy* **102**(6): 1200–1227.

Jovanovic B, Nyarko Y. 1995. A Bayesian learning model fitted to a variety of empirical learning curves. *Brookings Papers on Economic Activity Microeconomics* **1995**: 247–305.

Kaplan SN, Zingales L. 1997. Do investment-cash flow sensitivities provide useful measures of financing constraints? *Quarterly Journal of Economics* **112**(1): 169–215.

Levinthal DA. 1997. Adaptation on rugged landscapes. *Management Science* **43**(7): 934–950.

Levinthal DA, March JG. 1993. The myopia of learning. *Strategic Management Journal*, Winter Special Issue **14**: 95–112.

Levitt B, March JG. 1988. Organizational learning. *Annual Review of Sociology* **14**: 319–340.

Lieberman MB. 1984. The learning curve and pricing in the chemical processing industries. *Rand Journal of Economics* **15**(2): 213–228.

Lieberman MB. 1987. The learning curve, diffusion, and competitive strategy. *Strategic Management Journal* **8**(5): 441–452.

Malerba F. 1992. Learning by firms and incremental technical change. *Economic Journal* **102**(413): 845–859.

Muth JF. 1986. Search theory and the manufacturing progress function. *Management Science* **32**(8): 948–962.

Nelson RR, Winter SG. 1982. *An Evolutionary Theory of Economic Change.* Harvard University Press: Cambridge, MA.

Olley GS, Pakes A. 1996. The dynamics of productivity in the telecommunication equipment industry. *Econometrica* **64**(6): 1263–1297.

Pisano GP, Bohmer RMJ, Edmondson AC. 2001. Organizational differences in rates of learning: evidence from the adoption of minimally invasive cardiac surgery. *Management Science* **47**(6): 752–768.

Rapping L. 1965. Learning and World War II production functions. *Review of Economics and Statistics* **47**(1): 81–86.

Rivkin JW. 1998. Reconcilable differences: the relationship between industry conditions and firm effects. Working paper 98-065, Harvard Business School, Boston, MA.

Schmalensee R. 1989. Inter-industry studies of structure and performance. In *Handbook of Industrial Organization* (Vol. 2), Schmalensee R, Willig R (eds). Elsevier Science: Amsterdam, The Netherlands; 952–1009.

Simon HA. 1962. The architecture of complexity. *Proceedings of the American Philosophical Society* **106**(6): 467–482.

Sinclair G, Klepper S, Cohen W. 2000. What's experience got to do with it? Sources of cost reduction in a large specialty chemicals producer. *Management Science* **46**(1): 28–45.

Spence AM. 1981. The learning curve and competition. *Bell Journal of Economics* **12**(1): 49–70.

Thornton RA, Thompson P. 2001. Learning from experience and learning from others: an exploration of learning and spillovers in wartime shipbuilding. *American Economic Review* **91**(5): 1350–1368.

DOI: 10.1002/smj

APPENDIX A: INDUSTRY-BY-INDUSTRY LEARNING COEFFICIENTS

SIC	Rank	OLS	
	(OLS)	Coeff.	Std error
201	99	0.148	(0.02)
202	79	0.180	(0.02)
203	18	0.321	(0.02)
204	30	0.273	(0.02)
205	23	0.294	(0.02)
206	58	0.207	(0.04)
207	17	0.323	(0.04)
208	21	0.300	(0.02)
209	38	0.260	(0.02)
221	101	0.145	(0.07)
222	77	0.181	(0.04)
224	98	0.151	(0.08)
225	72	0.187	(0.02)
226	94	0.163	(0.03)
227	16	0.326	(0.05)
228	116	0.041	(0.03)
229	55	0.213	(0.03)
231	26	0.284	(0.04)
232	92	0.165	(0.02)
233	29	0.278	(0.01)
234	46	0.241	(0.04)
235	32	0.271	(0.08)
236	40	0.254	(0.03)
238	35	0.266	(0.04)
239	25	0.286	(0.02)
241	56	0.213	(0.01)
242	64	0.202	(0.01)
243	61	0.205	(0.01)
244	110	0.093	(0.02)
245	109	0.108	(0.02)
249	82	0.177	(0.02)
251	87	0.170	(0.01)
252	67	0.198	(0.03)
253	76	0.182	(0.04)
254	96	0.154	(0.02)
259	53	0.215	(0.03)
262	115	0.054	(0.04)
265	69	0.192	(0.01)
267	37	0.260	(0.03)
281	20	0.310	(0.02)
282	9	0.385	(0.03)
283	2	0.547	(0.03)
284	10	0.372	(0.03)
285	15	0.342	(0.03)
286	8	0.388	(0.03)
287	6	0.413	(0.03)
289	5	0.413	(0.02)
291	3	0.451	(0.10)
295	52	0.224	(0.02)
299	11	0.369	(0.05)
305	19	0.313	(0.05)
306	88	0.169	(0.02)
308	48	0.237	(0.01)
311	117	0.005	(0.07)
313	39	0.256	(0.09)

APPENDIX A: (*Continued*)

SIC	Rank	OLS	
	(OLS)	Coeff.	Std error
314	43	0.246	(0.03)
316	103	0.144	(0.08)
317	113	0.076	(0.07)
319	104	0.127	(0.07)
321	78	0.181	(0.05)
322	114	0.067	(0.03)
323	73	0.185	(0.02)
324	7	0.389	(0.08)
325	12	0.360	(0.03)
326	97	0.152	(0.03)
327	36	0.265	(0.01)
328	31	0.272	(0.04)
329	14	0.354	(0.02)
331	90	0.169	(0.02)
332	85	0.172	(0.02)
333	45	0.241	(0.05)
334	54	0.214	(0.05)
335	68	0.194	(0.02)
336	33	0.269	(0.03)
339	106	0.123	(0.03)
341	22	0.299	(0.02)
342	75	0.184	(0.02)
343	59	0.207	(0.03)
344	80	0.178	(0.01)
345	74	0.184	(0.02)
346	84	0.173	(0.02)
347	63	0.202	(0.01)
348	83	0.174	(0.05)
349	71	0.189	(0.01)
351	86	0.170	(0.04)
353	89	0.169	(0.02)
354	62	0.202	(0.01)
355	51	0.227	(0.02)
356	47	0.237	(0.01)
357	1	0.602	(0.03)
358	44	0.246	(0.02)
359	93	0.163	(0.01)
361	24	0.290	(0.02)
362	66	0.200	(0.02)
363	105	0.126	(0.04)
364	49	0.233	(0.02)
365	107	0.122	(0.04)
366	50	0.229	(0.03)
367	13	0.355	(0.02)
369	28	0.278	(0.02)
371	70	0.191	(0.01)
372	95	0.157	(0.02)
373	108	0.108	(0.02)
374	102	0.144	(0.05)
375	91	0.166	(0.06)
379	112	0.091	(0.02)
381	27	0.280	(0.07)
382	81	0.178	(0.02)
384	34	0.267	(0.02)
385	111	0.093	(0.04)
386	4	0.416	(0.04)

DOI: 10.1002/smj

APPENDIX A: (*Continued*)

SIC	Rank	OLS	
	(OLS)	Coeff.	Std error
391	42	0.250	(0.03)
393	100	0.146	(0.04)
394	41	0.251	(0.02)
395	65	0.202	(0.05)
396	57	0.209	(0.04)
399	60	0.206	(0.02)
Mean		0.227	
Std. dev		0.097	
Min		0.005	
Max		0.602	

APPENDIX B: ROBUSTNESS CHECKS

Are we measuring learning-by-doing?

We tested various other phenomena that might manifest as a high coefficient on cumulative output. While we cannot rule out all possible alternatives, some potentially important ones are addressed below.

Survivor bias and sample selection: With OLS, endogeneity of exit may bias the measured coefficient on cumulative output. If accumulated experience helps firms withstand bad 'performance shocks' and thus reduces the probability of exit, then the measured coefficient will be biased downward. On the other hand, if accumulated experience has no effect on exit, then the bias may be upward, the argument being that only 'good' firms survive and they would tend to have both higher cumulative output and higher productivity, resulting in a high 'learning' coefficient. However, these arguments do not appear to hold in our study for several reasons. If this argument were true, we should see a strong positive correlation between the turnover rate of firms (entry rate + exit rate) and our measured 'learning.' However, we see no statistically significant relationship between industry turnover rate and the measured learning rate (results available on request). Also, the Olley and Pakes (1996) and Ackerberg *et al*. (2006) estimates that correct for these potential biases are highly correlated with estimates presented here (results available on request). Further, the results suggest that learning is positively correlated with the cross-sectional heterogeneity of firm performance within an industry. If high measured learning rates reflected more selection, then one should observe a negative association between the measured learning rates and the cross-sectional heterogeneity of firm performance (as more firms are selected out in industries with high measured learning).

We also tested the robustness of our results to the choice of our sample. Using a sample that includes only ASM plants (which have better quality data) and relaxing our condition that plants not have a gap of more than two years between two consecutive years, produced learning estimates that were highly correlated with our baseline estimates (results available on request).

R&D investments: Sinclair *et al*. (2000) argue that it is specific R&D efforts that cause learning-by-doing. If all the learning were due to R&D, we should observe no learning once we include R&D as a control. Without controlling for R&D, we would observe high R&D industries to have high learning. Since the industries with high measured learning in our study are R&D intensive industries, we attempt to rule out R&D as *solely* driving the results. However, we lack detailed data on plant-level R&D and hence, we use firm-level R&D from Compustat as controls and reestimate the learning coefficients. The rank correlation between this set of learning coefficients and our original estimates is 0.67 and statistically significant at the 0.01 percent level. Furthermore, using these revised learning coefficients did not change the subsequent results on firm performance heterogeneity (Row 16 in Table 5).

Measurement errors in capital: It is well known that there are errors with measuring capital. If it were true that such measurement errors were more prevalent in some industries (e.g., in high-technology industries), then we may observe a high measured rate of learning in such industries. While there is no known way to completely rule this out, we reestimate our learning coefficients using an alternative measure of capital instead of the perpetual inventory method used in the study. Specifically, we use the year-end book value of assets and find that the resulting learning coefficients are highly correlated with our original estimates.

Industry life cycles: The need for learning-by-doing may be intricately linked to industry life cycles. For instance, early in the industry life cycle, firms may need to learn mostly on their own. Furthermore, this is also a period of great uncertainty and consequently higher heterogeneity of performance. As the industry matures, dominant design(s) emerge and firms may be able to

DOI: 10.1002/smj

benefit from others, thereby reducing the need for own learning. Concomitantly, the uncertainty also decreases, reducing performance heterogeneity. While this is certainly consistent with our arguments (note that we do not make any exhaustive claims over what gives rise to learning-by-doing), that we find exit rates to be uncorrelated with learning intensities suggests that industry life cycles are not the sole driving factor here.

Alternative production functions: We tested the robustness to relaxing the Cobb-Douglas production function form adopted here. Specifically, we estimated using OLS, a version of the translog production function $(y_{ijt} = a_{jt} + \alpha_j.k_{ijt} + \alpha'_j.(k_{ijt})^2 + \beta_j.l_{ijt} + \beta'_j.(l_{ijt})^2 + \gamma_j(k_{ijt})(l_{ijt}) + \lambda_j.x_{ijt} + \varepsilon_{ijt})$. These learning coefficients were highly correlated with our baseline estimates.

Learning-by-doing and heterogeneity of firm performance

In order to check the robustness of these results, we used the same type of regressions as in Table 4, but with different measures of performance heterogeneity, level of aggregation, choice of time periods, assumptions about error correlation structures, and others. Table 5 presents the results of robustness checks on this aspect. Each line in Table 5 is comparable to line 1 from Table 4.

Alternative measures of heterogeneity: So far, we have used the difference between the 90th and 10th percentiles as the measure of firm performance heterogeneity. In Table 5, Rows 1 and 2 use the interquartile range (difference between the 75th and 25th percentiles) and standard deviation of firm performance respectively. We observe that the coefficient on learning intensity is strongly positive.

Ordinal measures of learning intensity: In our baseline results, we directly used the estimated learning intensities as measures of learning-by-doing. Since the estimated learning intensities vary across different specifications, we check if using ordinal measures makes a difference. Row 3 of Table 5 uses the industry learning ranks (based on the estimated learning coefficient) instead of the estimated learning intensities. Row 4 uses a simple dummy variable that divides industries into high learning and low learning based on the median estimated learning coefficients. Once again, the results are statistically significant.

Excluding diversified firms: Our baseline results use the firm's primary SIC code to assign firms to industries. However, many firms in Compustat are diversified and hence, it may be argued that it is inappropriate to use a single learning coefficient for such firms. Using the Compustat business segments data from 1984 to 2000, we select firms that have a single three-digit SIC segment that comprises at least 95 percent of their total sales and estimate Equation 11 for these firms. Row 5 of Table 5 presents the results, which are strongly positive and statistically significant.

Level of aggregation: Since it may be argued that the three-digit SIC level is a very high level of aggregation, we estimate learning intensity at the four-digit SIC level and repeat our test. Row 6 of Table 5 shows that even with a finer industry definition, our results hold. Another possible concern is that the definitions of some three-digit SIC industries are simply much more heterogeneous than others. To rule this possibility out, (a) we excluded all three-digit SIC industries ending with '9' (usually 'not elsewhere classified' industries) and (b) included the number of four-digit SIC codes within a three-digit SIC as a control in Equation 11. The results remained statistically significant (Row 7, Table 5).

Other econometric concerns: Rows 8 to 16 of Table 5 provide results to alleviate other possible econometric concerns such as errors being correlated within an industry (robustness check: cluster errors within an industry), taking logarithms (robustness checks: not taking logarithms at all and not taking logarithms for the dependent variable), the choice of 10 firms in an industry year as the cutoff for inclusion (robustness check: use 25 firms as cutoff), influence of outliers (robustness check: exclude the two top and bottom industries in terms of learning rates, choice of time period (robustness check: separate 1973–1984 and 1985–2000), and other unobserved industry factors (robustness check: *joint* two-digit SIC year dummies). The results remain statistically significant for these alternative specifications.